DAMTP-2006-17
hep-th/0602233

The QCD Axion and Moduli Stabilisation

Joseph P. Conlon[1]

DAMTP, Centre for Mathematical Sciences, Wilberforce Road, Cambridge, CB3 0WA, UK.

Abstract

We investigate the conditions for a QCD axion to coexist with stabilised moduli in string compactifications. We show how the simplest approaches to moduli stabilisation give unacceptably large masses to the axions. We observe that solving the F-term equations is insufficient for realistic moduli stabilisation and give a no-go theorem on supersymmetric moduli stabilisation with unfixed axions applicable to all string compactifications and relevant to much current work. We demonstrate how nonsupersymmetric moduli stabilisation with unfixed axions can be realised. We finally outline how to stabilise the moduli such that f_a is within the allowed window $10^9\text{GeV} < f_a < 10^{12}\text{GeV}$, with $f_a \sim \sqrt{M_{SUSY} M_P}$.

[1]email:j.p.conlon@damtp.cam.ac.uk

Contents

1 Introduction

String theory remains certainly the most popular and on all appearances the best candidate for an ultraviolet completion of the Standard Model that will unify gauge and gravitational interactions in a consistent quantum theory. Since the discovery of Calabi-Yau compactifications [1], it has been hoped that the measured properties of the Standard Model may be understood through the details of a string compactification.

This has traditionally been approached top-down, starting with a fully specified and globally consistent closed string compactification. The defined technical problem is to construct a compactification whose low energy gauge group and matter content is that of the Standard Model or a close extension. The original context of this is the weakly coupled heterotic string on a Calabi-Yau, where the matter spectrum is determined by the geometry of the manifold and the gauge bundles thereon. Realistic matter content requires substantial mathematical effort; for an account of some recent developments [2] can be consulted.

The style of top-down constructions changed with the discovery of D-branes and the realisation that Standard-like models could also be constructed within 'intersecting brane worlds'. In these scenarios the Standard Model is located on

brane stacks that wrap cycles in the internal space, with the matter content determined by the intersection numbers of these cycles. Thus it is not just heterotic, but also (orientifolded) type II compactifications that can give pseudo-realistic physics. Such constructions have been developed both for toroidal orbifolds and smooth Calabi-Yaus, with a recent review being [3].

A general problem in top-down constructions is that of moduli stabilisation. The simplest string compactifications have very many moduli - massless scalar fields coupling with gravitational strength and determining the matter coupling constants. Such massless particles are however inconsistent with experiment and it is necessary to generate a potential for them. The difficulty is that moduli are associated with the geometry of the compactification, and the effects - preeminently fluxes - that stabilise the moduli also back-react on the geometry, making the resulting space difficult to study.

Together with the brane world picture, this has led to 'bottom-up' approaches to string phenomenology [4]. As branes are localised, the field theory on them depends only on local geometry. A bottom-up physicist first builds a local Standard Model, and only later worries about the global embedding. The limiting case of this is 'branes at singularities', where the entire low energy spectrum is determined by the nature of a pointlike singularity. In this context there have been recent attempts to find a Standard Model singularity [5]. A characteristic of this approach is that particle physics becomes an open string theory decoupled from the closed string dynamics associated with the compact geometry.

Regardless of the philosophy, any realistic string compactification must eventually account for the structure of the Standard Model. In this respect moduli stabilisation connects top-down and bottom-up constructions, as the moduli vevs determine the field theory coupling constants. One Standard Model coupling constant in particular need of explanation is the QCD θ angle. This is of course a well-known problem with a well-known answer: a Peccei-Quinn axion [6]. There do exist other possibilities, reviewed in [7], but for this paper I shall simply assume the Peccei-Quinn solution to be correct.

In the context of string theory, this creates a modulus anti-stabilisation problem. There are many string theory axions that may in principle solve the strong CP problem. To do so, an axion must remain massless throughout the thicket of moduli stabilisation effects and down to the QCD scale. This problem has previously been brought up in [8, 9]. It is clean and sharply posed, as effects very weak on the Planck scale may be very large on the QCD scale. Given the necessity of moduli stabilisation, this question is best analysed within the context of string constructions for which all moduli have been stabilised. The purpose of this paper is to determine for such constructions the conditions under which a QCD axion will exist.

The paper is organised as follows. I first review in section 2 the strong CP problem and the ways axions can arise in different string theory constructions. Section 3 investigates how to stabilise moduli while keeping axions sufficiently

light to solve the strong CP problem. I review some approaches to stabilising all moduli and show why their simplest versions contain no light axions. I investigate potentials with massless axions and show that in such cases there exist no supersymmetric minima of the potential, but nonsupersymmetric minima may exist. I give conditions on the geometry for a QCD axion to exist, both at leading and higher orders in the instanton expansion. Section 4 addresses the axion decay constant. In the context of the exponentially large volume compactifications of [10, 11] I outline how a phenomenologically acceptable value of f_a may be achieved together with the relationship $f_a \sim \sqrt{M_{SUSY} M_P}$.

2 Axions

2.1 Axions and the Strong CP Problem

In principle, the strong interactions can generate CP violation through an $F\tilde{F}$ coupling,

$$S_{F\tilde{F}} = \frac{\theta}{16\pi^2} \int F_{\mu\nu} \tilde{F}^{\mu\nu}. \tag{1}$$

However, observationally θ is extremely small: $|\theta| < 10^{-8}$. Given that at the above level, θ is a coupling constant which can *a priori* take values anywhere between 0 and 2π, this seems unnatural. The strong CP problem is to explain this observation.

There exist several proposed resolutions. In the context of string theory and string compactifications, the most natural is that due to Peccei and Quinn [6]. In this approach θ is promoted to a dynamical field, with Lagrangian

$$\mathcal{L} = \mathcal{L}_{SM} + \frac{1}{2} f_a^2 \partial_\mu \theta \partial^\mu \theta + \frac{\theta}{16\pi^2} F_{\mu\nu} \tilde{F}^{\mu\nu}. \tag{2}$$

In (2) f_a has dimensions of mass and is known as the axion decay constant. Conventionally a scalar has mass dimension one, and so we redefine $a \equiv \theta f_a$. This gives

$$\mathcal{L} = \mathcal{L}_{SM} + \frac{1}{2} \partial_\mu a \partial^\mu a + \frac{a}{16\pi^2 f_a} F_{\mu\nu} \tilde{F}^{\mu\nu}. \tag{3}$$

In equation (3) there exists an anomalous global $U(1)$ symmetry, $a \to a + \epsilon$. This symmetry is violated by QCD instanton effects, which break it to a discrete subgroup. These generate a potential for a,

$$V_{\text{instanton}} \sim \Lambda_{QCD}^4 \left(1 - \cos\left(\frac{a}{f_a} \right) \right). \tag{4}$$

In the absence of other effects, this potential is minimised at $a = 0$, setting the θ-angle to zero dynamically. The mass scale for the axion a is

$$m_a \sim \frac{\Lambda_{QCD}^2}{f_a}. \tag{5}$$

(A more precise estimate replaces Λ_{QCD}^2 by $f_\pi m_\pi$, where $f_\pi \sim 90\text{MeV}$ is the pion decay constant.) The Peccei-Quinn symmetry is by its nature anomalous. For the axion to solve the strong CP problem, the leading anomalous contribution to the potential must be that of QCD instantons, otherwise the minimum will be at $\theta \neq 0$.

Phenomenologically, f_a has only a narrow window of allowed values. The smaller the value of f_a, the more strongly the axion couples to matter. The condition that supernovae cool by predominantly emitting energy through neutrinos (rather than axions) is equivalent to the constraint $f_a > 10^9\text{GeV}$. There is also a cosmological upper bound on f_a. The axion field presumably starts its cosmological evolution with θ at some arbitrary value between 0 and 2π. Once the Hubble scale is comparable to the axion mass, the axion field oscillates and the energy density stored in the axion field is diluted with the expansion of the universe. The energy stored today thus depends on the axion mass, which is determined by f_a. The requirement that axions do not overclose the universe leads to the constraint $f_a < 10^{12}\text{GeV}$, with such axions being potential dark matter candidates.

The requirement $f_a < 10^{12}\text{GeV}$ follows from the assumption of a standard cosmology. String compactifications generally have many moduli, which in supersymmetric scenarios typically have masses $m \sim 1\text{TeV}$. These can be long-lived, causing problems with nucleosynthesis and giving low reheat temperatures. In supersymmetric models, the axion always has a scalar partner (the saxion) with the associated cosmological problems. It has been argued (in particular see [12, 13]) that the cosmological problems with the scalars are more severe than those associated with the axion, and so the upper bound on f_a should not be taken seriously without an associated resolution of the cosmological moduli problems. [13] reports that models can be found in which both axion and saxion cosmological problems may be evaded with $f_a \sim 10^{15}\text{GeV}$.

The cosmological history of the universe before nucleosynthesis is not known. However we are going to assume a standard cosmology and take the upper bound on f_a seriously. This is mainly because low reheat temperatures ($T \sim 10\text{MeV}$) have a generic problem that at reheating the decaying modulus has an $\mathcal{O}(1)$ branching fraction to gauginos and thus overproduces LSPs [14, 15]. In addition, low reheat temperatures make it hard to produce the observed baryon asymmetry. This suggests the universe should be hot at the weak scale, allowing a standard WIMP annihilation calculation and the possibility of electroweak baryogenesis, but also imposing the standard bounds on f_a. Furthermore, this is not in obvious contradiction with the properties of the saxion - its matter coupling is also set by f_a (rather than M_P), and low values of $10^9\text{GeV} < f_a < 10^{12}\text{GeV}$ can easily give large reheating temperatures $T \sim 10^5\text{GeV}$.

2.2 Axions in String Theory

String compactifications generically contain fields a_i which have $a_i F\tilde{F}$ couplings and possess the anomalous global symmetry $a \to a + \epsilon$ featuring in the axionic solution to the strong CP problem. We first enumerate possible axions, before considering their relation to the physics of moduli stabilisation.

Axions in the Heterotic String

In heterotic compactifications, the axions are traditionally divided into the universal, or model-independent, axion and the model-dependent axions. The model-independent axion is the imaginary part of the dilaton multiplet, $S = e^{-2\phi}\mathcal{V} + ia$. It is the dual of the 2-form potential $B_{2,\mu\nu}$ arising from the NS-NS 2-form field: $da = e^{-2\phi} * dB_{\mu\nu}$. The dilaton superfield is the tree-level gauge kinetic function for all gauge groups:

$$\mathcal{L} \sim \mathrm{Re}(S) \int F^a_{\mu\nu} F^{a\mu\nu} + \mathrm{Im}(S) \int F^a_{\mu\nu} \tilde{F}^{a\mu\nu}.$$

Consequently in a realistic compactification there must always exist an $aF_{QCD}\tilde{F}_{QCD}$ coupling.

There are also the model independent axions, b_i, given by the imaginary parts of the Kähler moduli T_i. For a basis Σ_i of 2-cycles of the Calabi-Yau, $T_i = t_i + ib_i$, with $t_i = \int_{\Sigma_i} J$ the string frame volume of the cycle Σ_i and $b_i = \int_{\Sigma_i} B_2$. At tree level, these have no couplings to QCD. However, such a coupling may be generated through the one loop correction to the gauge kinetic function. For the $E_8 \times E_8$ heterotic string, this correction is

$$\begin{aligned} f_1 &= S + \beta_i T_i, && (6)\\ f_2 &= S - \beta_i T_i, && (7) \end{aligned}$$

where 1 and 2 refer to the first and second E_8 respectively. The β_i are determined by the gauge bundles on the compactification manifold X. For gauge bundles V_1 and V_2,

$$\beta_i = \frac{1}{8\pi^2} \int e_i \wedge \left(c_2(V_1) - c_2(V_2)\right), \tag{8}$$

where e_i is the 2-form associated with the cycle Σ_i. This can be derived by dimensional reduction of the Green-Schwarz term $\int B_2 \wedge X_8(F_1, F_2, \mathcal{R})$.[2] The axions associated to the Kähler moduli are called model-dependent as their appearance in f depends on the one-loop correction, which in turn depends on the specific properties of the compactification.

[2] Strictly, this only gives the correction to $\mathrm{Im}(f)$. The corresponding correction to $\mathrm{Re}(f)$ is however implied by holomorphy.

Axions in Intersecting Brane Worlds

The discovery of D-branes [16] led to the extension of string model building beyond the heterotic string. The type II string theories, or more properly orientifolds thereof, can give rise to 'intersecting brane worlds'. In these the standard model is localised on a stack of branes while gravity propagates in the bulk. Light bifundamental matter arises from strings located at intersection loci and stretching between brane stacks.

The bosonic action of a single Dp-brane is

$$S_p = \frac{-2\pi}{(2\pi\sqrt{\alpha'})^{p+1}} \left(\int_\Sigma d^{p+1}\xi e^{-\phi} \sqrt{\det(g + B + 2\pi\alpha' F)} + i \int_\Sigma e^{B+2\pi\alpha' F} \wedge \sum_q C_q \right). \tag{9}$$

Σ is the cycle wrapped by the brane and the sum is a formal sum over all RR potentials in which only relevant terms are picked out.

The kinetic term $F_{\mu\nu}F^{\mu\nu}$ comes from the DBI action and the instanton action $F \wedge F$ from the Chern-Simons term. The gauge coupling corresponds to the inverse volume of Σ and the θ angle to the component of C_{p-3} along Σ. These fields pair up to become the scalar component of the chiral multiplet which is the gauge kinetic function of the resulting gauge theory.

As IIB compactifications are our main focus, we shall be more explicit here. In principle, IIB string theory allows, consistent with supersymmetry, space-filling D3, D5, D7 and D9-branes. However, in an orientifold setting we are restricted to either D3/D7 or D5/D9 pairs. We shall interest ourselves in the former case. The branes must wrap appropriate cycles to cancel the negative charge and tension carried by the orientifold planes; we assume this has been done.

In D3/D7 compactifications, the relevant superfields are those of the dilaton and Kähler moduli multiplets. Their scalar components are[3]

$$S = e^{-\phi} + ic_0, \tag{10}$$
$$T_i = \tau_i + ic_i. \tag{11}$$

c_0 is the Ramond-Ramond 0-form and $e^\phi \equiv g_s$ the string coupling. For Σ_i a 4-cycle of the Calabi-Yau,

$$c_i = \frac{1}{l_s^4} \int_{\Sigma_i} C_4 \qquad \text{and} \qquad \tau_i = \int_{\Sigma_i} \frac{e^{-\phi}}{2} J \wedge J,$$

where $l_s = 2\pi\sqrt{\alpha'}$ denotes the string length. Note that the Kähler modulus T_i involves the Einstein, rather than string, frame volume of a 4-cycle. This is

[3]Technically, this is only for the case that $h_-^{1,1} = 0$. This will not be important for the issues we discuss, and so we will use this simplifying assumption. The correct expressions under more general circumstances can be found in [17, 18].

most simply understood as the requirement that the gauge kinetic function be holomorphic in the chiral superfields. Indeed, S is the universal gauge kinetic function for D3-branes, whereas T_i is the gauge kinetic function for the field theory on a D7-brane stack wrapping the 4-cycle Σ_i.

The axionic couplings arise from the Chern-Simons term in the action. For D3-branes, this gives

$$S_{F\tilde{F}} = \frac{c_0}{4\pi} \int F \wedge F, \tag{12}$$

while for D7-branes

$$S_{F\tilde{F}} = \frac{c_i}{4\pi} \int F \wedge F. \tag{13}$$

By expanding the DBI action, we obtain the field theory couplings

$$\left.\frac{1}{g^2}\right|_{D3} = \frac{e^{-\phi}}{2\pi} \qquad \text{and} \qquad \left.\frac{1}{g^2}\right|_{D7} = \frac{\tau_i}{2\pi}.$$

There exists a similar story for IIA intersecting brane worlds, where the Standard Model must be realised on wrapped D6-branes (a Calabi-Yau has no 1- or 5-cycles to wrap D4- or D8-branes on). The gauge coupling now comes from the calibration form $\mathrm{Re}(\Omega)$ and the axion from the reduction of the 3-form potential C_3,

$$\left.\frac{1}{g^2}\right|_{D6} = \int_{\Sigma_i} \mathrm{Re}(\Omega) \qquad \text{and} \qquad \left.\theta\right|_{D6} = \int_{\Sigma_i} C_3.$$

Our interest is in the interaction of axions with moduli stabilisation and supersymmetry breaking, to which we now turn.

3 Axions and Moduli Stabilisation

It is obvious from the above that potential axions are easily found in string compactifications; indeed, they are superabundant. For axions to solve the strong CP problem, they must also be light, with QCD instantons giving the dominant contribution to their potential. Light axions are not in themselves problematic. Pure type II Calabi-Yau compactifications have many axions, which remain exactly massless as a consequence of four-dimensional $\mathcal{N} = 2$ supersymmetry. However, the same $\mathcal{N} = 2$ supersymmetry that guarantees the axions remain massless also guarantees a non-chiral spectrum with the axions' scalar partners massless. These are modes of the graviton and will lead to unobserved fifth forces.

More realistic string constructions have $\mathcal{N} = 1$ supersymmetry in four dimensions, allowing a potential to be generated for the moduli. To avoid bounds from fifth force experiments, the size moduli - the saxions - must at a minimum receive masses at a scale $m_T \gtrsim (100\mu\text{m})^{-1}$. However, the expected scale is much

larger: typical constructions give moduli masses comparable to the supersymmetry breaking scale, and the cosmological moduli problem [19, 20] suggests that in fact $m_T \gtrsim 10\text{TeV}$. In recent years there has been much progress in moduli stabilisation [21, 22], with physics such as fluxes and instantons used to lift the degeneracies associated with the geometric moduli. In the context of strong CP, this same progress creates a modulus anti-stabilisation problem. There are many stringy effects that can generate a potential for a putative QCD axion. These include worldsheet instantons, D-instantons and gaugino condensation, all of which are often invoked to stabilise the geometric moduli present. If any one of these effects is more important for a given axion than QCD instantons, that axion does not solve the strong CP problem.

Axion potentials generally come from nonperturbative effects whose magnitude is exponentially sensitive to the values of the stabilised moduli. The analysis of such effects is therefore best performed in a context within which all moduli are stabilised, and we first review mechanisms to achieve this. The most developed scenarios are those within IIB flux compactifications. We will work mostly with these, although we will along the way obtain a no-go theorem applicable to all string compactifications. IIB compactifications have the advantage that the stabilisation of dilaton and complex structure moduli can be studied ten-dimensionally [21], while the back-reaction of the fluxes is relatively mild: the internal geometry simply becomes conformally Calabi-Yau.

3.1 Review of Moduli Stabilisation

We shall orient our subsequent discussion around two approaches to stabilising all moduli, the well-known KKLT scenario [22] and the exponentially large volume models of [10, 11]. The point of departure is the flux compactifications of Giddings, Kachru and Polchinski [21]. These are compactifications of a IIB orientifold[4] in the presence of localised sources (D3/D7 branes and O3/O7 planes) with non-vanishing 3-form flux $G_3 = F_3 - SH_3$. Here F_3 and H_3 are RR and NSNS 3-form fluxes with S the dilaton-axion. The fluxes generate a superpotential[23]

$$W = \frac{1}{l_s^2} \int G_3 \wedge \Omega, \tag{14}$$

with Ω the unique holomorphic (3,0) form and the Kähler potential given by

$$\frac{\mathcal{K}}{M_P^2} = -2\ln(\mathcal{V}) - \ln\left(\int \Omega \wedge \bar{\Omega}\right) - \ln(S + \bar{S}). \tag{15}$$

This Kähler potential is no-scale for the Kähler moduli. The dilaton and complex structure moduli appear in the superpotential and are stabilised at leading order

[4]or more generally F-theory.

in the g_s and α' expansions. Their masses scale as

$$m_S \sim m_\phi \sim \frac{N}{R^6} M_P, \tag{16}$$

where N is a measure of the number of units of 3-form flux and R is the Calabi-Yau radius in units of l_s.

At this level, the no-scale property implies the Kähler moduli remain unstabilised. The dilaton and complex structure moduli are integrated out to focus on an effective theory for the Kähler moduli. Although absent at tree level, the Kähler moduli can appear nonperturbatively in the superpotential through brane instantons [24] or gaugino condensation,

$$W = W_0 + \sum_i A_i e^{-a_i T_i}, \tag{17}$$

where $a_i = 2\pi (\frac{2\pi}{N})$ for brane instantons (gaugino condensation). The KKLT proposal is to stabilise the Kähler moduli by solving $D_{T_i} W = 0$ for all i. The resulting 4-cycle sizes are

$$\tau_i \sim \frac{1}{a_i} \ln \left(\frac{W_0}{A_i} \right). \tag{18}$$

The cycle size goes logarathimically with W_0, and so to trust the supergravity approximation W_0 must be extremely small. We may usefully rephrase this condition. The KKLT construction relies on nonperturbative corrections to the scalar potential dominating the perturbative corrections. This is equivalent to the requirement that the 'correction' to the superpotential is comparable to, or larger, than the tree-level term. In order for this to hold at (moderately) large T, W_0 must be very small, which is to be achieved by tuning fluxes.

However, for almost all values of W_0 ('almost' can be made precise as in [25]) there exists no reliable regime in which this requirement is met. Perturbative corrections, originating from corrections to the Kähler potential, are then essential for the study of the moduli potential. It was shown in [10, 11] that the incorporation of an α'^3 correction to the Kähler potential [26],

$$\mathcal{K} = -2 \ln(\mathcal{V}) \rightarrow \mathcal{K} = -2 \ln \left(\mathcal{V} + \frac{\xi}{2 g_s^{\frac{3}{2}}} \right), \tag{19}$$

combined with the same nonperturbative superpotential corrections

$$W = W_0 + \sum_{i=2}^{h^{1,1}} A_i e^{-a_i T_i},$$

leads, subject to a necessary condition $h^{2,1} > h^{1,1} > 1$, to a minimum of the potential at exponentially large volumes:

$$\mathcal{V} \sim W_0 e^{\frac{c}{g_s}},$$

for a model-dependent constant c. This and other properties of the minimum follow from an explicit study of the scalar potential, but a summary of the results is as follows:

1. The moduli divide into one large modulus τ_b and $h^{1,1}-1$ small 'blowup' moduli τ_i. Whereas the former is exponentially large, $\tau_b \sim \mathcal{V}^{\frac{2}{3}}$, the latter all have $\tau_i \sim \ln(\mathcal{V})$. The origin of this is as follows. If we take $\tau_b^{\frac{3}{2}} \sim \mathcal{V} \gg 1$ and fixed, the other moduli τ_i minimise their potential at $D_{T_i}W = 0 + \mathcal{O}\left(\frac{1}{\mathcal{V}}\right)$. Neglecting numerical factors, this generates an effective potential for $\mathcal{V}$ of

$$V \sim -\frac{(\ln \mathcal{V})^{\frac{3}{2}}}{\mathcal{V}^3} + \frac{\xi}{g_s^{\frac{3}{2}}\mathcal{V}^3},$$

where the $\frac{\xi}{\mathcal{V}^3}$ comes from the Kähler correction. $\mathcal{V}$ is then dynamically stabilised at $\mathcal{V} \gg 1$.

2. The stabilised volume is exponentially sensitive to the stabilised string coupling, allowing a natural generation of hierarchies: the structure of the potential is such as to create a hierarchical separation of the string and Planck scales.

The Kähler correction (19) - of course only one of many - plays an essential role in this construction. The effects of other possible Kähler corrections, such as from warping effects or other IIB $\mathcal{O}(\alpha'^3)$ corrections, have been extensively discussed in [11] (see also [28]). We will comment further on these in the appendix. Here we simply note that the only hope of incorporating Kähler corrections in a controlled manner is via a small expansion parameter. In this case $\mathcal{V} \gg 1$ and the expansion parameter is $\frac{1}{\mathcal{V}}$. Of known corrections, (19) gives the leading contribution to the scalar potential in the $\frac{1}{\mathcal{V}}$ expansion.

Finally let us note that the two constructions above both give rise to AdS minima, which are respectively supersymmetric and non-supersymetric. It is necessary to include a phenomenological uplift term, arising from e.g. a $\overline{\text{D3}}$-brane in a warped throat, in order to match the observed cosmological constant. In the KKLT scenario, it is the uplift that gives rise to soft supersymmetry breaking terms. For the exponentially large volume compactifications, the supersymmetry breaking is inherited from the no-scale structure and the effects of the uplift are subdominant [11, 29].

3.2 The Simplest Scenarios: Why Axions are Heavy

We now turn to the axions. Our first point is that the simplest versions of the above scenarios lack a QCD axion. All potential axions receive a high scale mass and thus cannot solve the strong CP problem. For simplicity we concentrate on

the KKLT construction, but a very similar argument holds for the exponentially large volume compactifications.

We start by asking whether QCD is to be realised on D3 or D7 branes. If we were to use D3-branes, the QCD axion would be the imaginary component of the dilaton multiplet, $S = e^{-\phi} + ic_0$. However, as indicated in (16) this multiplet is stabilised at tree-level by the fluxes, with a mass $m_S \sim \frac{N}{R^6} M_P$, with R the radius in units of l_s. This tree-level stabilisation may seem at odds with the axionic shift symmetry $c_0 \to c_0 + 2\pi$. However, the shift symmetry is a subgroup of the fundamental $SL(2,Z)$ symmetry, under which the fluxes also transform. By nature a duality, this is invisible in the low-energy theory and cannot protect the axion from acquiring a mass. As $R \lesssim 5$ in KKLT, the axion obtains a mass $m_{c_0} \gtrsim 10^{15}$GeV and cannot be a QCD axion. For the exponentially large volume compactifications, R is larger but the conclusion unchanged: QCD on a D3 brane stack is inconsistent with the existence of a Peccei-Quinn axion.

This implies that QCD ought to be realised on D7 branes. The axions are now the imaginary parts of the Kähler moduli, and the instanton effects used to stabilise these moduli will also give the axions a mass. To estimate the scale of this mass, it is simplest just to construct the potential explicitly.

As above, we take the superpotential

$$W = W_0 + \sum_{i=1}^{h^{1,1}} A_i \exp(-a_i T_i), \tag{20}$$

with Kähler potential

$$\mathcal{K} = -2 \log \mathcal{V}. \tag{21}$$

Note that $\mathcal{K} = \mathcal{K}(T_i + \bar{T}_i)$, and so the axions do not appear either in $\mathcal{K}$ or its derivatives. The supergravity F-term potential is

$$V = e^{\mathcal{K}} (\mathcal{K}^{i\bar{j}} D_i W D_{\bar{j}} \bar{W} - 3|W|^2). \tag{22}$$

The no-scale property of the Kähler potential simplifies (22) to

$$V = e^{\mathcal{K}} \left(\mathcal{K}^{i\bar{j}} \partial_i W \partial_{\bar{j}} \bar{W} + \mathcal{K}^{i\bar{j}} \left((\partial_i \mathcal{K}) W \partial_{\bar{j}} \bar{W} + (\partial_{\bar{i}} \mathcal{K}) \bar{W} \partial_j W \right) \right). \tag{23}$$

It is a property of the Kähler potential (21) that $\mathcal{K}^{i\bar{j}} \partial_i \mathcal{K} = -2\tau_j$. (23) becomes

$$\begin{aligned} V \;=\; & e^{\mathcal{K}} \Bigg(\mathcal{K}^{i\bar{j}} a_i a_j \left(A_i \bar{A}_j e^{-a_i T_i - a_j \bar{T}_j} + \bar{A}_i A_j e^{-a_i \bar{T}_i - a_j \bar{T}_j} \right) \\ & -2 a_i \tau_i \left(W \bar{A}_i e^{-a_i \bar{T}_i} + \bar{W} A_i e^{-a_i T_i} \right) \Bigg). \end{aligned} \tag{24}$$

It is easy to extract the axionic dependence of the potential (24). $\mathcal{K}$ and its derivatives are all real and phases only come from the superpotential. The potential

becomes

$$\begin{aligned} V &= e^{\mathcal{K}}\Big(\mathcal{K}^{i\bar{j}}\left(2a_ia_j|A_iA_j|e^{-a_i\tau_i-a_j\tau_j}\cos(a_i\theta_i+a_j\theta_j+\gamma_{ij})\right) \\ &\quad -4a_i\tau_i|W_0A_i|e^{-a_i\tau_i}\cos(a_i\theta_i+\beta_i)-4a_i\tau_i|A_iA_j|\cos(a_i\theta_i+a_j\theta_j+\gamma_{ij})\Big). \end{aligned} \tag{25}$$

θ_i denote the axions and the phases γ_{ij} and β_i come from the phases of $A_i\bar{A}_j$ and $\bar{A}_iW$ respectively. The axionic mass matrix is

$$M_{ij}^2 = \frac{\partial^2 V}{\partial\theta_i\partial\theta_j}, \tag{26}$$

and we obtain[5]

$$M_{ij}^2 \sim a_ia_jV_{min}, \tag{27}$$

with V_{min} the magnitude of the potential at the AdS minimum. In KKLT, $D_{T_i}W = 0$ for all i and $V_{min} \sim -3\frac{|W_0|^2}{\mathcal{V}^2}$. As there are $h^{1,1}$ independent phases, there is no reason for M_{ij}^2 to be degenerate and we expect all eigenvalues to be $\mathcal{O}(a^2V_{min})$, where a is the typical magnitude of the a_i.

The determination of physical masses also requires the Kähler potential. In general there is no explicit expression for the overall volume $\mathcal{V}$ in terms of 4-cycle volumes τ_i. The Kähler metric may however be written as [17]

$$\mathcal{K}_{i\bar{j}} = \frac{G^{-1}_{i\bar{j}}}{\mathcal{V}^2}, \qquad G_{i\bar{j}} = -\frac{3}{2}\left(\frac{k_{ijk}v^k}{\mathcal{V}} - \frac{3}{2}\frac{k_{imn}t^mt^nk_{jpq}t^pt^q}{\mathcal{V}^2}\right). \tag{28}$$

If $\mathcal{V} \sim$ (a few)l_s^6, $\mathcal{K}_{i\bar{j}} \sim \mathcal{O}(1)$ and the mass matrix M_{ij}^2 gives a good estimate of the scale of axion masses. If $\mathcal{V} \gg l_s^6$, then $\mathcal{K}_{i\bar{j}} \ll \mathcal{O}(1)$, and M_{ij}^2 underestimates the axion masses. M_{ij}^2 could only overestimate the axion masses if $\mathcal{V} \ll 1$. This realm of moduli space is not accessible in a controlled fashion and we do not concern ourselves with it.

In units where $M_P = 1$, we therefore have

$$m_{\tau_i} \sim m_{c_i} \sim a_i\sqrt{V_{min}} \sim \frac{a_iW_0}{\mathcal{V}}. \tag{29}$$

The axion masses are consequently set by the value of the tree-level superpotential W_0. This also determines the vacuum energy and, implicitly, the energy scale of supersymmetry breaking required to cancel the vacuum energy. TeV-scale soft terms require hierarchically small W_0. For the (gravity-mediated) case studied in [30], this required $W_0 \sim 10^{-13}$, with

$$m_{\tau_i} \sim m_{a_i} \sim m_{3/2} \sim 10\text{TeV}. \tag{30}$$

[5]We emphasise this does not mean the mass matrix is $M^2 = a \otimes a$: simply that terms may receive an enhance by factors appearing in the exponential.

This scale is vastly greater than that associated with QCD instanton effects, and thus the axions are incapable of solving the strong CP problem. We could insist on a QCD axion, and require that W_0 be sufficiently small that QCD instantons dominate over the D-instanton effects of moduli stabilisation. This would require $W_0 \sim 10^{-40}$. However, this scenario is entirely excluded as the size moduli are light enough to violate fifth force experiments and the susy breaking scale would be $\mathcal{O}(10^{-14}\text{eV})$. Consequently, in the simplest KKLT scenario it is impossible to generate a QCD axion. The D3 axion receives a high scale mass from fluxes, whereas the instanton effects give the D7 axions large masses comparable to the size moduli.

A similar argument holds for the exponentially large volume compactifications. As in KKLT, the D3 axion receives a large flux-induced mass. The 'small' cycles are stabilised by instanton effects, and these give the corresponding axions masses of a similar scale to the size moduli, $m_{a_i} \sim m_{\tau_i} \sim m_{3/2}$. One difference is that there is a modulus, the 'large' modulus τ_b, which need not appear in the superpotential. It is stabilised through the Kähler potential, and while it is massive its axionic partner indeed remains massless. However, this cycle is exponentially large, and any gauge group supported on this cycle is far too weakly coupled to be QCD. The same conclusion holds: the simplest version of this scenario does not generate a QCD axion.

The above formulates the 'modulus anti-stabilisation problem': naive scenarios of moduli stabilisation are incompatible with a QCD axion.

We next examine an apparent solution to this problem that in fact has a subtle flaw. We want a way to stabilise moduli without stabilising the axions. Axions correspond to phases in the superpotential and do not appear in the Kähler potential. If we included a multi-exponential term $e^{-\alpha^i T_i}$ in the superpotential, a massless axion would certainly survive, as at least one phase would be absent. As the size moduli all appear in the Kähler potential, by solving the F-term equations we may hope to stabilise the size moduli while leaving the axions massless.

To illustrate this idea, let us consider a toy model,

$$\begin{aligned} \mathcal{K} &= -\ln(T_1+\bar{T}_1) - \ln(T_2+\bar{T}_2) - \ln(T_3+\bar{T}_3), && (31)\\ W &= W_0 + Ae^{-2\pi(T_1+T_2+T_3)}. && (32) \end{aligned}$$

The Kähler potential (31) is that appropriate for toroidal orbifolds, with $\mathcal{V} = t_1 t_2 t_3$. To understand where the superpotential could arise from, we can hypothesise that the cycle (1+2+3) is the smallest cycle with only two fermionic zero modes, and that instantons wrapping (for example) the cycle (2+3) all have more than two zero modes and do not appear in the superpotential. However, we are not here really concerned with the microscopic origin of the superpotential: at this level we simply regard equations (31) and (32) as defining the model.

The F-term equations $D_{T_1}W = D_{T_2}W = D_{T_3}W = 0$ give

$$-2\pi Ae^{-2\pi(T_1+T_2+T_3)} - \frac{1}{T_1+\bar{T}_1}\left(W_0 + Ae^{-2\pi(T_1+T_2+T_3)}\right) = 0, \qquad (33)$$

$$-2\pi A e^{-2\pi(T_1+T_2+T_3)} - \frac{1}{T_2+\bar{T}_2}(W_0 + Ae^{-2\pi(T_1+T_2+T_3)}) = 0, \quad (34)$$

$$-2\pi A e^{-2\pi(T_1+T_2+T_3)} - \frac{1}{T_3+\bar{T}_3}(W_0 + Ae^{-2\pi(T_1+T_2+T_3)}) = 0. \quad (35)$$

These immediately imply $\tau_1 = \tau_2 = \tau_3$ and equations (33) to (35) collapse to

$$2\pi A e^{-6\pi\tau_1} e^{-2\pi i(\theta_1+\theta_2+\theta_3)} + \frac{1}{2\tau_1}(W_0 + Ae^{-6\pi\tau_1}e^{-2\pi i(\theta_1+\theta_2+\theta_3)}) = 0. \quad (36)$$

While the sum $\theta_1+\theta_2+\theta_3$ is fixed, there are clearly two axionic directions not relevant for the solution of the F-term equations. On the other hand, there is a unique value for the size moduli such that the F-term equations are solved. Except for the massless axionic directions, the scales of the masses are unaltered from above, and we would expect

$$m_{\tau_i} \sim m_{\theta_1+\theta_2+\theta_3} \sim \frac{W_0}{\mathcal{V}}, \qquad m_{\theta_1-\theta_2} = m_{\theta_1-\theta_3} = 0. \quad (37)$$

As this is supergravity rather than rigid supersymmetry, there is no contradiction in having a mass splitting for the multiplet in the presence of unbroken supersymmetry.

While this seems promising, there is in fact a serious problem with the above. Even though all F-term equations can be solved, numerical investigation shows that at the supersymmetric locus the resulting scalar potential is tachyonic, with signature $(+,-,-)$. Although supersymmetry ensures the moduli are Breitenlohner-Freedman stable [31], this notion of AdS stability ceases to be relevant after the (necessary) uplift.

We now show that these tachyons are in fact generic for any attempt to stabilise the moduli supersymmetrically while preserving unfixed axions.

3.3 A No-Go Theorem

We suppose we have an arbitrary $\mathcal{N}=1$ supergravity with moduli fields, Φ_α, $T_\beta = \tau_\beta + ib_\beta$, where the b_β are the axions. We write the superpotential and Kähler potential as

$$W = W(\Phi_\alpha, T_\beta), \quad (38)$$

$$\mathcal{K} = \mathcal{K}(\Phi_\alpha, T_\beta + \bar{T}_\beta). \quad (39)$$

The Peccei-Quinn symmetry $b_\beta \to b_\beta + \epsilon_\beta$ implies the form of (39) should hold in perturbation theory.

We further suppose we have solved

$$D_{\Phi_\alpha} W = 0 \text{ and } \qquad D_{T_\beta} W = 0 \quad (40)$$

for all α and β, but that at least one axion $b_u = \sum_\beta \lambda_\beta b_\beta$ is unfixed: the solution to (40) is independent of $\langle b_u \rangle$.

We redefine the basis of chiral superfields so that there exists a superfield T_u with $b_u = \mathrm{Im}(T_u)$,

$$\begin{aligned} T_1 &\rightarrow T_u, \\ T_2 &\rightarrow T_2, \\ T_n &\rightarrow T_n. \end{aligned} \tag{41}$$

This is a good redefinition as it does not affect holomorphy properties.

As the solution to all F-term equations is independent of b_u, b_u is a flat direction of the potential (22) at the supersymmetric locus.[6] The potential at the supersymmetric locus is given by

$$V = -3e^{\mathcal{K}}|W|^2. \tag{42}$$

As b_u does not appear in $\mathcal{K}$, it follows that if b_u is a flat direction $|W|$ must be independent of b_u. Up to one exception this then implies that W is independent of b_u.

The sole exception is if b_u purely represents an overall phase, i.e. $W = e^{-aT_u}$ with no constant term. In a IIB context, this may potentially arise if the flux superpotential exactly vanishes due to a discrete symmetry [32], while a combination of non-perturbative effects and Kähler corrections stabilise the Kähler moduli. While potentially interesting, this is an exceptional case and we do not analyse it further.

If W has no explicit dependence on $b_u = \mathrm{Im}(T_u)$, it follows by holomorphy that it also has no explicit dependence on $\tau_u = \mathrm{Re}(T_u)$ and hence on T_u. Therefore

$$\partial_{T_u} W \equiv 0. \tag{43}$$

However, as $D_{T_u} W = 0$, it follows that at the supersymmetric locus,

$$\text{either} \qquad (\partial_{T_u} K) = 0 \qquad \text{or} \qquad W = 0.$$

The latter is overdetermined and non-generic, so we first focus on $(\partial_{T_u} K) = 0$.

Direct calculation now shows that the τ_u direction is tachyonic at the supersymmetric locus. To see this, note that from the scalar potential (22)

$$\partial_{\tau_u} V = e^{\mathcal{K}} \mathcal{K}^{i\bar{j}} \left(\partial_{\tau_u}(D_i W) D_{\bar{j}} \bar{W} + D_i W \partial_{\tau_u}(D_{\bar{j}} \bar{W}) \right) - 3(\partial_{\tau_u} K) e^K W \bar{W}. \tag{44}$$

[6]The requirement of *flatness* is stronger than the requirement that the axion simply be *massless*. Flatness is the right requirement, as if an axion is fixed in any way it does not solve strong CP.

We have used $\partial_{\tau_u} W \equiv 0$ and have only kept terms that will give non-vanishing contributions to $\partial_{\tau_u}\partial_{\tau_u} V$ at the supersymmetric locus. Expanding $D_i W$ and again using $\partial_{\tau_u} W \equiv 0$, (44) simplifies to

$$\partial_{\tau_u} V = e^{\mathcal{K}}\mathcal{K}^{i\bar{j}} \left(\partial_{\tau_u}(\partial_i \mathcal{K}) W (D_{\bar{j}}\bar{W}) + D_i W \partial_{\tau_u}(\partial_{\bar{j}}\mathcal{K})\bar{W}\right) - 3(\partial_{\tau_u}\mathcal{K}) e^{\mathcal{K}} W\bar{W}. \quad (45)$$

If we again only keep terms non-vanishing at the supersymmetric locus, the second derivative is

$$\partial_{\tau_u}\partial_{\tau_u} V = e^{\mathcal{K}}\mathcal{K}^{i\bar{j}} \left(2\partial_{\tau_u}(\partial_i \mathcal{K})\partial_{\tau_u}(\partial_{\bar{j}}\mathcal{K}) W\bar{W}\right) - 3(\partial_{\tau_u}\partial_{\tau_u}\mathcal{K}) e^{\mathcal{K}} W\bar{W}. \quad (46)$$

Now, as $\tau_u = \frac{1}{2}(T_u + \bar{T}_u)$,

$$\partial_{\tau_u}\mathcal{K}(T+\bar{T}) = 2\partial_{T_u}\mathcal{K}(T+\bar{T}),$$

and we have

$$\begin{aligned} \partial_{\tau_u}\partial_{\tau_u} V &= 4e^{\mathcal{K}} W\bar{W}(2\mathcal{K}^{i\bar{j}}\mathcal{K}_{iu}\mathcal{K}_{u\bar{j}} - 3\mathcal{K}_{u\bar{u}}) \\ &= -4e^{\mathcal{K}} W\bar{W}\mathcal{K}_{u\bar{u}}, \end{aligned} \quad (47)$$

where we have used $\mathcal{K}_{i\bar{j}} = \mathcal{K}_{ij}$. As $\mathcal{K}_{i\bar{j}}$ is a metric, $\mathcal{K}_{u\bar{u}}$ is positive definite and it follows that the τ_u direction is tachyonic.

Now consider the $W = 0$ case. As indicated above, this is non-generic: even if W originally vanishes, it is expected to receive non-perturbative corrections which make it non-vanishing. Even so, it follows easily that

$$\partial_{\tau_u}\partial_{\tau_u} V = 0, \quad (48)$$

and so the τ_u size modulus is massless, leading to unobserved fifth forces.

The above gives a no-go theorem: there does not exist any supersymmetric minimum of the F-term potential consistent with stabilised moduli and unfixed axions.

It is in the nature of no-go theorems that they admit loopholes, so let us discuss ways around this result. One point to consider is the form of the Kähler potential, as we have used in the above argument the fact that

$$\mathcal{K} = \mathcal{K}(T+\bar{T}). \quad (49)$$

While true in perturbation theory because of the axionic shift symmetry, this equation will break down nonperturbatively, and the argument showing that the τ_u direction is tachyonic will no longer hold. However, this same breakdown will cause $\mathcal{K}$, and hence the potential V, to depend on the axion. As this lifts the required axionic flat direction, the no-go theorem will cease to apply.

A second loophole is that although the F-term potential might be tachyonic, the D-term potential might come to the rescue. For example, a Fayet-Iliopoulos

term might have exactly the right structure to render the supersymmetric locus an actual minimum of the full potential. However this seems implausible in the presence of many tachyonic directions. A similar approach would be to try and set $W = 0$ and then rely entirely on D-terms to stabilise the moduli. Another possibility (discussed recently in [33]) is that an anomalous U(1) might remove the tachyonic directions. While the massive gauge boson will eat the axionic degree of freedom, an axionic direction may survive in the phase of a scalar charged under the U(1).

A third loophole is that stability might not rely on the existence of an actual minimum for the potential. Equation (47) involves the Kähler metric $\mathcal{K}_{u\bar{u}}$. The kinetic term for τ_u is $\mathcal{K}_{u\bar{u}}\partial_\mu \tau_u \partial^\mu \tau_u$. If we just consider the τ_u direction, the physical mass is therefore

$$m^2_{\tau_u} = -2e^{\mathcal{K}} W\bar{W} = -\frac{8}{9}|m_{BF}|^2, \tag{50}$$

where m_{BF} is the relevant Breitenlohner-Freedman bound. As tachyonic modes can be stable in AdS, one could argue that it is sufficient simply to solve the F-term equations and not to worry about whether the resulting locus is an actual minimum of the potential.

While this point is more substantial, it does not resolve the problem. The real world is not AdS, and for stability requires a positive definite mass matrix. For any realistic model, the vacuum energy must be uplifted such that it vanishes. After this uplift, the extra geometric advantages of AdS go away and the tachyons can no longer be supported. As there may be many tachyons present - one for each massless axion - the entire problem of moduli stabilisation must necessarily be solved over again in the uplifting. As the uplift is generally the least controlled part of the procedure, this seems a hard problem. While the uplift *may* remove the tachyons, at the present level of understanding this seems pure hypothesis. It is then very unclear how useful the original supersymmetric AdS saddle point is, and whether it is a suitable locus to uplift.

The argument above suggests that supersymmetric solutions are unpromising starting points from which to address the strong CP problem. Either all moduli appear in the superpotential, in which case there is no light axion, or a modulus is absent from the superpotential, in which case the potential is tachyonic. The fourth and most obvious loophole is then to give up on the requirement of supersymmetric minima, and search for nonsupersymmetric minima of the potential with massless axions.

We shall consider this point in the next section. Before we do so we also observe that while our focus here is on axions, the argument above is also relevant for moduli stabilisation. For example, in the weakly coupled heterotic string, the one-loop corrections to the gauge kinetic function (6) imply that gaugino condensation generates a superpotential

$$W_{n.p.} = Ae^{-\alpha S + \beta_i T^i}. \tag{51}$$

It has been proposed to use the superpotential (51), together with a constant term W_0, to stabilise the dilaton and Kähler moduli by solving $D_S W = D_{T^i} W = 0$. However, there is only one phase - and hence only one axion - explicitly present in (51). The above argument shows that the resulting scalar potential will actually be tachyonic at the supersymmetric locus, with signature $(+, -, \ldots, -)$.

A similar result will apply to the recent study of IIA flux compactifications with all NSNS and RR fluxes turned on. In this context it is also found that the solution of the F-term equations is independent of many of the axions present (those associated with the C_3 field). The above implies that as long as $h^{2,1} \neq 0$ the supersymmetric locus is always tachyonic, with one tachyon present for every massless axion. Tachyons have been recognised in particular models [33, 34], but from the above they would seem to be very generic.

3.4 Non-supersymmetric Minima with Massless Axions

The above no-go theorem shows that supersymmetric moduli stabilisation is not a good starting point from which to solve the strong CP problem. This implies we ought to consider non-supersymmetric moduli stabilisation. That there is no no-go theorem for non-supersymmetric minima with massless axions can be shown by construction: for example, the exponentially large volume compactifications of [10, 11] all contain a massless axion associated with the large cycle controlling the overall volume. As indicated above, this cannot be a QCD axion, as any brane on this cycle is very weakly coupled. If we want to try and force this cycle into being a QCD axion, we can tune the parameters to force the minimum of this potential to relatively small volumes. That this is possible can be confirmed numerically. Another possibility would be a purely perturbative stabilisation of the volume modulus, solely using Kähler corrections (in which axions do not appear). This has been discussed in [27, 28], although without an explicit example.

We shall not dwell on these possiblities. First, because the resulting axion decay constant would be, as we shall see in section 4, close to the Planck scale and outside the allowed window and secondly, because at such small volumes there is no good control parameter. As we require non-supersymmetric minima we shall base our discussion around the large-volume models of [10, 11]. For now we only discuss moduli stabilisation but in section 4 we shall show that these can also realise phenomenological values for f_a.

We illustrate the discussion with a three-modulus toy model, in which we assume the volume may be expressed in terms of 4-cycles as

$$\mathcal{V} = (T_1 + \bar{T}_1)^{\frac{3}{2}} - (T_2 + \bar{T}_2)^{\frac{3}{2}} - (T_3 + \bar{T}_3)^{\frac{3}{2}}. \tag{52}$$

We use three moduli as this turns out to be the minimal number required for our purposes: clearly, this is no significant restriction. Expressed in terms of 2-cycles, (52) corresponds to

$$\mathcal{V} = \lambda(t_1^3 - t_2^3 - t_3^3). \tag{53}$$

We note (53) satisfies the requirement that $\frac{\partial^2 \mathcal{V}}{\partial t_i \partial t_j}$ have signature $(+,-,-)$. We may perhaps think of this toy model as a $\mathbb{P}^3$ with two points blown up. Denoting the cycles by 1, 2 and 3, the Kähler potential is

$$\mathcal{K} = -2\ln\left((T_1+\bar{T}_1)^{\frac{3}{2}} - (T_2+\bar{T}_2)^{\frac{3}{2}} - (T_3+\bar{T}_3)^{\frac{3}{2}}\right) - \frac{\xi}{g_s^{3/2}\mathcal{V}}, \tag{54}$$

where we have included the leading large-volume behaviour of the α'^3 correction of [26]. g_s is fixed by the fluxes and in (54) should be regarded as a tunable parameter. For superpotential, we shall take

$$W = W_0 + e^{-\frac{2\pi}{n}(T_2+T_3)}. \tag{55}$$

This could arise from gaugino condensation on a stack of n branes wrapping the combined cycle 2+3.[7] QCD will be realised as a stack of branes wrapping cycle 3.

In the limit $\mathcal{V} \gg 1$ with τ_2 and τ_3 small, the leading functional form of the scalar potential is (omitting numerical factors)

$$V = \frac{(\sqrt{\tau_2}+\sqrt{\tau_3})e^{-\frac{2\pi}{n}2(\tau_2+\tau_3)}}{\mathcal{V}} - \frac{(\tau_2+\tau_3)e^{-\frac{2\pi}{n}(\tau_2+\tau_3)}}{\mathcal{V}^2} + \frac{\xi}{g_s^{3/2}\mathcal{V}^3}. \tag{56}$$

The minus sign in (56) arises from minimising the potential for the axion $\mathrm{Im}(T_2+T_3)$. The axions $\mathrm{Im}(T_1)$ and $\mathrm{Im}(T_2-T_3)$ do not appear in (56) and are unfixed. By considering the limit $\mathcal{V}\to\infty$, $\frac{2\pi(\tau_2+\tau_3)}{n} \sim \ln\mathcal{V}$, it follows that as $\mathcal{V}\to\infty$ the potential (56) goes to zero from below. As by adjusting g_s we can make the third term of (56) arbitrarily large, we can ensure the potential remains positive until arbitrarily large volumes, and thus any minimum will be at exponentially large volumes.

Is there a minimum? The potential is clearly symmetric under $\tau_2 \leftrightarrow \tau_3$, and the potential restricted to the locus $\tau_2 = \tau_3$ indeed has a minimum at exponentially large volumes. Because of the symmetry $\tau_2 \leftrightarrow \tau_3$, this 'minimum' is also a critical point of the full potential. However, it is not a minimum of the full potential. At fixed $\tau_2+\tau_3$ and fixed $\mathcal{V}$, (56) depends only on $\sqrt{\tau_2}+\sqrt{\tau_3}$. For fixed $\tau_2+\tau_3$, this is *maximised* at $\tau_2=\tau_3$, and the mode $\tau_2-\tau_3$ is tachyonic at this locus. We have not investigated whether this tachyon satisfies the Breitenlohner-Freedman bound for AdS stability. This is for the same reasons as above: once we uplift, the geometric protection of AdS ceases to be relevant. Consequently the fields in (56) run away either to $\tau_2 = 0$ or $\tau_3 = 0$, where one of the blow-up cycles collapses.

This result shows that the above toy model does not, by itself, have a minimum of the potential with a massless QCD axion. We may ask whether this is a feature

[7] The use of gaugino condensation rather than instanton effects is necessary to ensure that the cycle 2+3 is large enough to contain QCD.

of the geometric details of the model - for example, whether a different choice of triple intersection form in (53) would alter this result. We have investigated several other toy models without finding a minimum, and while we have no proof we suspect none exists so long as the Kähler potential is given by (54).

This is bad news, but it is controllable bad news. The instability above is very particular: there is no instability either for the overall volume or for the sum of the blow-up volumes $\tau_1 + \tau_2$, but only for the difference $\tau_1 - \tau_2$. The effect of the instability is to drive one of the blow-up cycles to collapse. Consequently, the instability can be cured by *any* effect that becomes important at small cycle volume and prevents collapse.

For example, the addition of a term

$$\frac{1}{\sqrt{\tau_2}\mathcal{V}^3} + \frac{1}{\sqrt{\tau_3}\mathcal{V}^3} \tag{57}$$

to (56) would obviously stabilise the cycles τ_2 and τ_3 against collapse and generate a minimum of the potential. As this term does not affect the argument that in the $\mathcal{V} \to \infty$ limit the potential approaches zero from below, the resulting minimum would still be at exponentially large volume.

We discuss in greater detail in the appendix which Kähler corrections are and are not allowed. For now we note that terms of the form (57) may be generated from a correction to the Kähler potential,

$$\mathcal{K} + \delta\mathcal{K} = -2\ln(\mathcal{V}) + \frac{\epsilon\sqrt{\tau_2}}{\mathcal{V}} + \frac{\epsilon\sqrt{\tau_3}}{\mathcal{V}}. \tag{58}$$

For simplicity we have kept the $\tau_2 \leftrightarrow \tau_3$ symmetry. Such a correction is motivated by the fact that it gives corrections to the Kähler metrics $\mathcal{K}_{2\bar{2}}$ and $\mathcal{K}_{3\bar{3}}$ suppressed by factors of g^2 for the field theory on the relevant cycle. More specifically,

$$\mathcal{K}_{2\bar{2}} + \delta\mathcal{K}_{2\bar{2}} = \frac{3}{2\sqrt{2\tau_2}\mathcal{V}}\left(1 - \frac{\epsilon}{12\sqrt{2}\tau_2}\right). \tag{59}$$

As $\tau_2 = \frac{1}{g^2}$ for a brane wrapping the cycle 2, the correction is suppressed by g^2.

The inverse metric involves an infinite series of terms diverging in the $\tau_2 \to 0$ and $\tau_3 \to 0$ limit. For example,

$$\mathcal{K}^{2\bar{2}} = \frac{2\sqrt{2\tau_2}\mathcal{V}}{3}\left(1 + \frac{\epsilon}{12\sqrt{2}\tau_2} + \frac{\epsilon^2}{288\tau_2^2} + \dots\right). \tag{60}$$

This is easy to understand: at $\tau_2 = \frac{\epsilon}{12\sqrt{2}}$, the Kähler metric $\mathcal{K}_{2\bar{2}}$ goes to zero and the inverse metric diverges. This divergence can be seen by resumming (60). In the physical potential, this divergence will create a positive wall at finite values of τ_2 and τ_3. The positivity can be seen from the fact that the divergence in $\mathcal{K}^{-1}$ will only appear in the term

$$e^{\mathcal{K}}\mathcal{K}^{i\bar{j}}D_iW D_{\bar{j}}\bar{W}, \tag{61}$$

which is manifestly positive definite. Consequently the potential will diverge positively at finite values of τ_2 and τ_3, and so a stable minimum must exist for both τ_2 and τ_3.

We have now outlined, in the context of the scenario of [10, 11], a way to stabilise all the size moduli while keeping a massless QCD axion left over. We would like cycle 3 to support QCD: by adjusting ϵ, we can always make the correction (58) sufficiently large to ensure that τ_3 is stabilised with the correct size for QCD. For intermediate string scales, this requires $\tau_3 \sim 10$. As the actual correction would be very hard to calculate, at this level we just adjust ϵ phenomenologically. Of course, the complexity of a real model is much greater than that of (58). However we note again that, even though the corrections cannot be calculated, our proposal for moduli stabilisation only requires that they exist and come with the right sign to prevent collapse.

While the above has been with a toy example, the above approach will apply to any model in which the moduli are stabilised along the lines of [10, 11]. Keeping an axion massless introduces an instability causing a blow-up cycle to want to collapse. Kähler corrections that become important at small volume can stabilise this cycle but will not affect the overall structure of the potential, and in particular will not affect the stabilisation of the volume at $\mathcal{V} \gg 1$.

3.5 Higher Instanton Effects in the Axion Potential

We have given above a Kähler potential and superpotential that will stabilise the moduli while containing a candidate QCD axion. The nonperturbative terms in the superpotential are in general just the leading terms in an instanton expansion. Even though the higher order terms may be highly suppressed and irrelevant to moduli stabilisation, they could still lift the flat direction associated with the massless axion. Given that $\Lambda_{QCD} \ll M_P$, even highly subleading terms could dominate over QCD instantons.

Let us estimate the general magnitude of such instanton effects. The magnitude of brane instantons depends on the volumes of the cycles they can wrap. Generally there will be many such cycles, whose sizes depend on the stabilised moduli, but minimally there must always exist the cycle which support the QCD stack. It follows from the DBI action that the gauge coupling for a D7-brane stack is[8]

$$\frac{1}{g^2} = \frac{\mathrm{Re}(T)}{4\pi} \Rightarrow \alpha^{-1} = \frac{4\pi}{g^2} = \mathrm{Re}(T) = \tau.$$

This defines the gauge coupling at the high scale where the effective field theory becomes valid. When $m_s \sim M_P$, this is in essence the string scale, but if $m_s \ll$

[8]v3: This corrects a factor of 2 from earlier versions. To deduce the gauge coupling from reduction of the DBI action, it is necessary to ensure the gauge group generators are normalised according to the phenomenology conventions $Tr(T^\alpha T^\beta) = \frac{1}{2}\delta^{\alpha\beta}$.

Table 1: Cycle sizes and instanton amplitudes for various UV scales

E_{UV}	10^8 GeV	10^{10} GeV	10^{12} GeV	10^{14} GeV	10^{16} GeV
$\alpha_{QCD}^{-1}(\mathrm{E}_{UV})$	15.8	18.1	20.4	22.7	25
$\mathrm{Re(T)} = \alpha^{-1}$	15.8	18.1	20.4	22.7	25
$e^{-2\pi T}$	7.7×10^{-44}	2×10^{-50}	2×10^{-56}	6×10^{-63}	6×10^{-69}
$e^{-4\pi T}$	5.9×10^{-87}	4×10^{-100}	4×10^{-112}	3.6×10^{-127}	3.6×10^{-139}

M_P, the difference between m_{KK} and m_s become significant. It is a subtle issue whether m_s or m_{KK} is the appropriate high scale. If QCD is supported on a small cycle within a large internal space, the KK modes associated with the bulk will be uncharged under QCD and will not contribute to the running coupling. KK modes of the QCD cycle will contribute, but these will be at masses comparable to the string scale. In considering the running coupling, we therefore use m_s as the high scale rather than m_{KK}. We consider a wide range of string scales and take a sampling of high scale values from $10^8 \to 10^{16}$ GeV. Given a string scale, the internal volume is determined by $m_s \sim \frac{M_P}{\sqrt{\mathcal{V}}}$.

The QCD coupling runs logarithmically with energy scale, with

$$\alpha_{QCD}^{-1}(10^2\mathrm{GeV}) \sim 9 \qquad \text{and} \qquad \alpha_{QCD}^{-1}(10^{16}\mathrm{GeV}) \sim 25.$$

The required high-scale couplings and cycle sizes are given in table 1, together with the action for a D3-brane instanton wrapping the same cycle as the QCD stack. Its magnitude is set by $\sim e^{-2\pi T}$ and we show in table 1 the approximate magnitude of single- and double-instanton effects. In addition to the QCD cycle, there may be other cycles which instantons may wrap. We do not include these for two reasons: first, whether such instantons would generate a potential for the QCD axion is model-dependent[9], and secondly, we can always arrange the model such that the QCD cycle is smallest and hence dominates the instanton expansion. We observe that the magnitude of the required cycle volume, and thus the magnitude of potential instanton effects, varies significantly with the string scale. If present, such D-instantons would generate a potential for the QCD axion. To compare their magnitude to that of QCD effects, we need an estimate of their contribution to the scalar potential. In this context we only care about terms containing a phase and so contributing to the axion potential. To this end, the relevant term from the scalar potential (23) is

$$V_{axion} = e^{\mathcal{K}} \left(\mathcal{K}^{i\bar{j}} \left(\partial_i W (\partial_{\bar{j}} \mathcal{K}) \bar{W} + c.c. \right) \right). \tag{62}$$

[9] It seems odd that such instantons could affect the QCD axion at all. However, if QCD is on cycle 3, and the axion $b_2 + b_3$ is fixed by effects on cycle 2+3, an instanton solely on cycle 2 effectively generates a potential for the QCD axion.

Table 2: Magnitude of Axion Potentials from Superpotential Instanton Effects

E_{UV}	10^8 GeV	10^{10} GeV	10^{12} GeV	10^{14} GeV	10^{16} GeV
$\mathcal{V}$	10^{20}	10^{16}	10^{12}	10^8	10^4
$V_{\text{1-instanton}}$	$10^{-79} M_P^4$	$10^{-81} M_P^4$	$10^{-83} M_P^4$	$10^{-85} M_P^4$	$10^{-87} M_P^4$
$V_{\text{2-instanton}}$	$10^{-122} M_P^4$	$10^{-131} M_P^4$	$10^{-139} M_P^4$	$10^{-151} M_P^4$	$10^{-155} M_P^4$

A superpotential instanton contribution $e^{-2\pi n T_i}$ generates a term

$$V_{axion} = \frac{-2a_i \tau_i W_0}{\mathcal{V}^2} e^{-2\pi n \tau_i} \cos(\theta_i). \tag{63}$$

The absolute magnitude of (63) depends on the value of n, the internal volume $\mathcal{V}$ and the tree-level superpotential W_0. As we are looking towards phenomenology we also take $m_{3/2} = \frac{W_0}{\mathcal{V}} \sim 1\text{TeV} \sim 10^{-15} M_P$, as appropriate for gravity-mediated TeV-scale soft terms. Note that for models built around the KKLT scenario, we always have $m_s \gtrsim M_{GUT}$ and only the largest value of E_{UV} is achievable. In table 2 we give the internal volumes required for each UV scale, as well as the resulting absolute magnitude of 1-,2- and 3-instanton superpotential corrections to the scalar potential. For the same reasons as above, we only consider instantons wrapping the QCD cycle.

As well as superpotential effects, there are also nonperturbative corrections to the Kähler potential (the perturbative corrections to $\mathcal{K}$ do not have an axionic dependence). While smaller, these are easier to generate - the instantons can have four fermionic zero modes rather than only two. A correction

$$\mathcal{K} = -2\ln(\mathcal{V}) \to \mathcal{K} = -2\ln(\mathcal{V} + e^{-2\pi n T}) \tag{64}$$

will generate effects in the scalar potential at order

$$V_{\delta\mathcal{K}} \sim \frac{W_0^2}{\mathcal{V}^3} e^{-2\pi n T}.$$

and thus generate a potential for the QCD axion θ of the form

$$V_{\delta K} \cos(\theta + \alpha).$$

Again assuming TeV-scale (visible) SUSY breaking, $\frac{W_0}{\mathcal{V}} \sim 10^{-15}$, the magnitudes of such effects are shown in table 3.

The axion potential originating from QCD effects and relevant to the strong CP problem is

$$V_{QCD} \sim \Lambda_{QCD}^4 (1 - \cos(\theta)),$$

with $\Lambda_{QCD} \sim 2 \times 10^{-19} M_P$ and $\Lambda_{QCD}^4 \sim 10^{-75} M_P^4$. We require QCD effects to be sufficiently dominant to be consistent with the failure to observe CP violation in strong interactions. Suppose we have a potential

$$V = A(1 - \cos(\theta)) + \epsilon \cos(\theta + \gamma). \tag{65}$$

Table 3: Magnitude of Axion Potentials from Kähler Potential Instanton Effects

E_{UV}	10^8 GeV	10^{10} GeV	10^{12} GeV	10^{14} GeV	10^{16} GeV
V	10^{20}	10^{16}	10^{12}	10^8	10^4
$V_{\text{1-instanton}}$	$10^{-99}M_P^4$	$10^{-97}M_P^4$	$10^{-95}M_P^4$	$10^{-93}M_P^4$	$10^{-91}M_P^4$
$V_{\text{2-instanton}}$	$10^{-143}M_P^4$	$10^{-147}M_P^4$	$10^{-151}M_P^4$	$10^{-155}M_P^4$	$10^{-159}M_P^4$

If $A \gg \epsilon$, the minimum is displaced from $\theta = 0$ by $\delta\theta \sim \frac{\epsilon}{A}$. Observationally, $|\theta| < 10^{-10}$, and thus non-QCD contributions must have absolute magnitude smaller than $10^{-85}M_P^4$. By comparison with tables 2 and 3 it follows that in order for QCD instantons to dominate the axion potential, the instanton corrections to the Kähler potential are not constrained, while the 1-instanton superpotential correction may or may not be present depending on factors of 2π and the precise value of the string scale. Higher instanton corrections to the superpotential are not constrained.

To contribute to a superpotential (Kähler potential) an instanton must have at most 2 (4) fermionic zero modes, to generate $\int d^4x d^2\theta$ and $\int d^4x d^2\theta d^2\bar{\theta}$ terms respectively. In the absence of flux, there is a necessary condition on a divisor to generate a superpotential [24]: it must have holomorphic Euler characteristic one, $\chi_g(D) = 1$. In the presence of flux, this condition may be relaxed. The number of zero modes on an instanton, and thus its ability to appear in either the Kähler or superpotential, may also be affected by the presence of the stack of QCD branes wrapping the would-be instanton cycle. We shall not attempt to analyse this question for specific 'real' models, but by fiat simply assume the necessary instantons to be absent from the potential. In this regard it is encouraging that the number of instantons required to be suppressed is quite limited.

4 The Axion Decay Constant

4.1 Magnitude

The last section was devoted to solving the strong CP problem: keeping an axion light while stabilising the moduli. However, even achieving this does not resolve all phenomenological problems. Given that a QCD axion exists, as indicated in the introduction there are strong bounds on the axion decay constant f_a of equation (3): $10^9\text{GeV} < f_a < 10^{12}\text{GeV}$. The lower bound, from supernova cooling, is hard. While the upper bound may be relaxed by considering non-standard cosmologies, here we shall also treat this as hard. We want to estimate f_a in some moduli stabilisation scenarios. In IIB compactifications, the axionic coupling to QCD arises from the clean and model-independent Chern-Simons coupling. How-

ever, to obtain the physical value of f_a the axion must be canonically normalised. This depends on the Kähler metric, and in particular on where the moduli are stabilised.

We assume we can write the Kähler potential as

$$\mathcal{K} = \mathcal{K}(T_i + \bar{T}_i), \tag{66}$$

with $\mathcal{K}$ real. This is true in perturbation theory, owing to the axionic shift symmetry, and any nonperturbative violations are small enough to be irrelevant for this purpose. In this case the kinetic terms for the axionic and size moduli do not mix. Noting that $\mathcal{K}_{i\bar{j}} = \mathcal{K}_{j\bar{i}}$, we have for any i and j

$$\begin{aligned} & \mathcal{K}_{i\bar{j}}(\partial_\mu T^i \partial^\mu \bar{T}^j) + \mathcal{K}_{j\bar{i}}(\partial_\mu T^j \partial^\mu \bar{T}^i) \\ = \ & \mathcal{K}_{i\bar{j}}\left((\partial_\mu \tau_i + i\partial_\mu c_i)(\partial^\mu \tau_j - i\partial^\mu c_j) + (\partial_\mu \tau_j + i\partial_\mu c_j)(\partial^\mu \tau_i - i\partial^\mu c_i)\right) \\ = \ & \mathcal{K}_{i\bar{j}}(2\partial_\mu \tau_i \partial^\mu \tau_j + 2\partial_\mu c_i \partial^\mu c_j), \end{aligned} \tag{67}$$

and the two sets of terms decouple.

Let us first show that if both the overall volume and the cycle volumes are comparable to the string scale, then as expected $f_a \gtrsim 10^{16}$GeV. Suppose an axion c_i is to be the QCD axion. The Lagrangian for this axion is

$$\mathcal{K}_{i\bar{i}}\partial_\mu c_i \partial^\mu c_i + \frac{c_i}{4\pi}\int F^a \wedge F^a. \tag{68}$$

For simplicity we have not included mixing terms: these will not greatly affect the discussion.

The simplest toy model is that of a factorisable toroidal orientifold, with Kähler potential

$$\begin{aligned} \mathcal{K} &= -\ln\left((T_1 + \bar{T}_1)(T_2 + \bar{T}_2)(T_3 + \bar{T}_3)\right) \\ &= -\ln(T_1 + \bar{T}_1) - \ln(T_2 + \bar{T}_2) - \ln(T_3 + \bar{T}_3). \end{aligned} \tag{69}$$

and Kähler metric

$$\mathcal{K}_{i\bar{j}} = \begin{pmatrix} (T_1 + \bar{T}_1)^{-2} & 0 & 0 \\ 0 & (T_2 + \bar{T}_2)^{-2} & 0 \\ 0 & 0 & (T_3 + \bar{T}_3)^{-2} \end{pmatrix}. \tag{70}$$

If we denote the axions by c_1, c_2 and c_3, the axion kinetic terms are

$$\frac{1}{4\tau_1^2}\partial_\mu c_1 \partial^\mu c_1 + \frac{1}{4\tau_2^2}\partial_\mu c_2 \partial^\mu c_2 + \frac{1}{4\tau_3^2}\partial_\mu c_3 \partial^\mu c_3. \tag{71}$$

For definiteness, let us assume QCD is realised on cycle 1. There is no inter-axion mixing and the relevant axion Lagrangian is

$$\frac{1}{4\tau_1^2}\partial_\mu c_1 \partial^\mu c_1 + \frac{c_1}{4\pi}\int F^a \wedge F^a. \tag{72}$$

If we canonically normalise $c_1' = \frac{c_1}{\sqrt{2}\tau_1}$, this becomes

$$\frac{1}{2}\partial_\mu c_1' \partial^\mu c_1' + \frac{\sqrt{2}\tau_1}{4\pi} c_1' \int F^a \wedge F^a. \qquad (73)$$

In units where $M_P = 1$, the axion decay constant is

$$f_a = \frac{1}{4\pi\tau_1\sqrt{2}}.$$

If QCD is to be realised on this cycle, we need $\tau_1 \sim 12$, and thus $f_a \sim 10^{16}$GeV. Going beyond this toy example, we recall that in general the Kähler metric was given by (28),

$$\mathcal{K}_{i\bar{j}} = \frac{G^{-1}_{i\bar{j}}}{\mathcal{V}^2}, \qquad G_{i\bar{j}} = -\frac{3}{2}\left(\frac{k_{ijk}t^k}{\mathcal{V}} - \frac{3}{2}\frac{k_{imn}t^m t^n k_{jpq}t^p t^q}{\mathcal{V}^2}\right). \qquad (74)$$

If all cycles are string scale in magnitude, then $\mathcal{K}_{i\bar{j}} \sim \mathcal{O}(1)$ and it is impossible to lower the axion decay constant substantially through canonical normalisation. The same conclusion applies: $f_a \gtrsim 10^{16}$GeV. This conclusion is unsuprising: the axionic coupling to matter is a stringy coupling, and so we expect f_a to be comparable to the string scale. If the string and Planck scales are identical, f_a cannot lie within the allowed window.

If we lower the string scale, phenomenological values for f_a can be achieved. To analyse this, let us return to the toy model of (52). We recall the Kähler potential was

$$\mathcal{K} = -2\ln\left((T_1 + \bar{T}_1)^{\frac{3}{2}} - (T_2 + \bar{T}_2)^{\frac{3}{2}} - (T_3 + \bar{T}_3)^{\frac{3}{2}}\right). \qquad (75)$$

The Kähler metric for this model is

$$\mathcal{K}_{i\bar{j}} = \begin{pmatrix} \frac{-3}{2\sqrt{2\tau_1}\mathcal{V}} + \frac{9\tau_1}{\mathcal{V}^2} & -\frac{9\sqrt{\tau_2}}{2\mathcal{V}^{5/3}} & -\frac{9\sqrt{\tau_3}}{2\mathcal{V}^{5/3}} \\ -\frac{9\sqrt{\tau_2}}{2\mathcal{V}^{5/3}} & \frac{3}{2\sqrt{2\tau_2}\mathcal{V}} + \frac{9\tau_2}{\mathcal{V}^2} & \frac{9\sqrt{\tau_2\tau_3}}{\mathcal{V}^2} \\ -\frac{9\sqrt{\tau_3}}{2\mathcal{V}^{5/3}} & \frac{9\sqrt{\tau_2\tau_3}}{\mathcal{V}^2} & \frac{3}{2\sqrt{2\tau_3}\mathcal{V}} + \frac{9\tau_3}{\mathcal{V}^2} \end{pmatrix}. \qquad (76)$$

The axion kinetic terms are $\mathcal{K}_{i\bar{j}}\partial_\mu c_i \partial^\mu c_j$. At small volumes there is substantial mixing between the axions c_1, c_2 and c_3. However, in the limit $\mathcal{V} \to \infty$ with $\tau_1 \gg \tau_2, \tau_3$, the Kähler metric takes the schematic form

$$\mathcal{K}_{i\bar{j}} \sim \begin{pmatrix} \mathcal{V}^{-4/3} & \mathcal{V}^{-5/3} & \mathcal{V}^{-5/3} \\ \mathcal{V}^{-5/3} & \mathcal{V}^{-1} & \mathcal{V}^{-2} \\ \mathcal{V}^{-5/3} & \mathcal{V}^{-2} & \mathcal{V}^{-1} \end{pmatrix}, \qquad (77)$$

and is to a good approximation diagonal. The requirement $\tau_1 \gg 1$ implies that QCD cannot be realised on branes wrapping cycle 1, as the resulting field theory

is far too weakly coupled. However, if $\tau_2 \sim \tau_3 \sim 10$ we may realise QCD by wrapping branes on one of these cycles. The resulting axion decay constant is

$$f_a \sim \frac{\sqrt{\mathcal{K}_{3\bar{3}}}}{4\pi} M_P \sim \frac{\mathcal{O}(1)}{4\pi\sqrt{\mathcal{V}}} M_P. \tag{78}$$

Thus if $\mathcal{V} \sim 10^{14}$ and $\tau_3 \sim 10$, the QCD gauge coupling is correct and the axion decay constant $f_a \sim 10^{10}\text{GeV}$ lies within the narrow phenomenological window. Up to $\mathcal{O}(1)$ factors, the string and Planck scales are related by

$$m_s = \frac{g_s M_P}{\sqrt{\mathcal{V}}}. \tag{79}$$

Such a large volume corresponds to lowering the string scale to $m_s \sim 10^{11}\text{GeV}$. The lowered axion decay constant is easy to understand physically. f_a measures the axion-matter coupling, which is an effect localised around the small QCD cycle. Thus the only scale it is sensitive to is the string scale, and so up to numerical factors $f_a \sim m_s$.

The above is a very particular limit of moduli space, with one cycle taken extremely large while all others are only marginally larger than the string scale. It would thus be essentially a curiosity if it were not also the exact regime in which the moduli are stabilised in the compactifications of [10, 11] reviewed above. As the stabilised volume is exponentially sensitive to the stabilised dilaton, *a priori* the string scale can lie anywhere between the Planck and TeV scales. There is no difficulty, and no fine-tuning, in stabilising the volume so as to achieve an intermediate string scale.

The above result on f_a is independent of whether an axion remains massless or not. As described in section 3.2, the simplest version of the scenarios of [10, 11] makes all axions far too heavy to solve the strong CP problem. In sections 3.4 and 3.5 we have described the necessary modifications to this scenario such that a massless QCD axion will survive to solve the strong CP problem. Combining this with the above, we have for the first time given a procedure to stabilise all moduli while ensuring a QCD axion exists with a phenomenologically allowed value for f_a.

In itself this is interesting, as axions within the phenomenological window have always been hard to achieve in string compactifications. However, this scenario compels a further very interesting relationship between the axion decay constant and the (visible) supersymmetry breaking scale.

4.2 Relation to Supersymmetry Breaking Scale

We argued earlier that if a QCD axion is to be present in IIB flux compactifications, QCD must be realised on a stack of D7-branes. We have also described

how to stabilise the moduli such that a QCD axion can exist with a phenomenologically allowed decay constant.

By definition, any scenario of moduli stabilisation determines the moduli vevs and masses. However, in general much more information can be extracted. If the moduli potential breaks supersymmetry, it will generate soft supersymmetry breaking terms in the visible sector. In gravity-mediated scenarios, it is these that dominate the soft MSSM Lagrangian. It is an old and important problem to go from the moduli potential to the soft terms. The analysis of this problem was initiated in the heterotic string. As full moduli stabilisation in that context is hard, the relevant moduli potential was unknown. Progress was nonetheless achieved by parametrising supersymmetry breaking as S,T or U-dominated, depending on the moduli multiplet (dilaton, Kähler or complex structure) in which the dominant F-term occurred. This allowed a classification of supersymmetry breaking possibilities even in the absence of a full moduli potential.

This question has now been revisited in IIB flux compactifications, where the moduli potential is under better control and full stabilisation is achievable (at least at the level of effective field theory). Instead of having to assume the structure of the F-terms, they may now be computed directly from the moduli potential. However, the problem is still subtle as the vacuum is originally AdS and the form of the (necessary) uplift to Minkowksi space can introduce extra contributions to the F-terms. In the KKLT scenario, the AdS minimum is supersymmetric and all details of the supersymmetry breaking therefore lie in the uplift. Unfortunately this is the part of the construction under least control. Using some particular assumptions about the uplift potential, (gravity-mediated) supersymmetry breaking in this scenario has been studied in [30].

For the exponentially large volume compactifications, the structure of supersymmetry breaking has been studied in [11, 29]. The original AdS minimum is non-supersymmetric, and the structure of the F-terms is essentially inherited from the no-scale potential of [21]. While the uplift is again not well controlled, fortunately its effects are subdominant. This arises because the magnitude of the vacuum energy in the AdS minimum is $\mathcal{O}\left(\frac{1}{\mathcal{V}^3}\right)$, whereas from the F-terms it 'ought' to be $\mathcal{O}\left(\frac{1}{\mathcal{V}^2}\right)$, the difference of course being accounted for by the no-scale cancellation. This implies that 'extra' F-terms required to lift the minimum to Minkowski space will be hierarchically smaller than those already present in the AdS minimum.

No-scale breaking corresponds to Kähler-domination, with the dominant F-term associated with the multiplet controlling the overall volume [11]. The structure of the soft terms on D7-branes for the no-scale models of [21] has been well-studied. Supersymmetry breaking is transmitted to the observable sector through gravitational interactions, and the 3-form fluxes present induce soft terms on the worldvolume theory of the D7 branes[36, 37]. Here we are only concerned with the results rather than the calculational details: the scalar and gaugino masses

are found to be

$$m_{D7} \sim M_{D7} \sim m_{3/2} = e^{\mathcal{K}/2} W \sim \frac{M_P}{\mathcal{V}}. \tag{80}$$

If the internal volume $\mathcal{V}$ is exponentially large, it is the prime determinant of the scale of the soft terms. Other, more model-dependent, factors will also be present, but these will only be $\mathcal{O}(1)$ effects relevant for the detailed structure but not the overall scale. As the volume also sets the scale of the axion decay constant, this implies a numerical relationship between these two quantities. From the fact that

$$f_a \sim \frac{M_P}{\sqrt{\mathcal{V}}} \qquad \text{and} \qquad m_{soft} \sim \frac{M_P}{\mathcal{V}},$$

it follows that up to numerical factors

$$f_a = \sqrt{M_P m_{soft}}. \tag{81}$$

Thus in such models the axion decay constant is compelled to be the geometric mean of the Planck scale and the (visible) supersymmetry breaking scale. This is a striking result, as the two pieces of physics are *a priori* entirely unrelated. The origin of this can be traced to the intermediate string scale, which has been argued to have interesting phenomenological properties[38].

5 Conclusions

The purpose of this paper has been to investigate the conditions under which a QCD axion, ideally with a phenomenological value for f_a, may coexist with stabilised moduli in string compactifications. This divides into two questions: first, how to stabilise the moduli such that a massless axion survives, and secondly, how to obtain allowed values of f_a, $10^9\text{GeV} < f_a < 10^{12}\text{GeV}$.

In the context of the first question, we have shown that the simplest version of many moduli stabilisation scenarios do not contain any light axions. We also have a negative result, in that supersymmetric moduli stabilisation is disfavoured: there exist no supersymmetric minima of the F-term potential with flat axionic directions. Even if AdS stability is present due to the Breitenlohner-Freedman bound, the tachyons must be removed by the time we are in Minkowski space. Performing this step requires a much greater technical understanding of uplifting AdS vacua to Minkowksi space than is currently available, and so it is unclear how relevant the original supersymmetric AdS solution is.

This result is pure $\mathcal{N} = 1$ supergravity and so makes no assumptions about the particular string model considered. It thus applies to all string compactifications, and in particular shows that in many of the supersymmetric IIA flux compactifications considered recently the complex structure moduli sector is heavily tachyonic.

There is no no-go theorem on nonsupersymmetric minima with massless axions, and so these may be preferred. In the context of the large volume compactifications of [10, 11], we outlined how to stabilise moduli while keeping axions massless. Here we had to rely on Kähler corrections that will become important as a cycle collapses to zero size. While unfortunately not much is known about these, our main requirement was simply that they exist. Clearly progress in determining the form of such corrections would be very interesting. We also specified the extent to which subleading higher-order instantons must be absent in order for a leading-order axion to solve the strong CP problem.

We note this result also favours gravity-mediated supersymmetry breaking. If the moduli potential must break supersymmetry in order to solve the strong CP problem, then this suggests that supersymmetry should be broken at the string scale. Gravity mediation therefore always contributes to the visible soft terms and, unless the string scale is lowered to $\mathcal{O}(1000\text{TeV})$, will dominate over gauge mediated effects. An intermediate string scale may then be preferred in order to obtain TeV-scale soft terms.

In the context of the second question, the fact that f_a is hierarchically lower than the Planck scale implies that compactifications with $m_s \sim M_P$ are unlikely to give allowed values for f_a. In the compactifications of [10, 11] the string scale is hierarchically lower than the Planck scale. In these models, $f_a \sim M_s$ and $M_{SUSY} \sim \frac{M_s^2}{M_P}$. An intermediate string scale therefore gives both $10^9\text{GeV} < f_a < 10^{12}\text{GeV}$ and visible susy breaking at $\mathcal{O}(1\text{TeV})$. It is hard to find models with phenomenological values for f_a, and so it is very interesting that in the above model this also implies TeV-scale supersymmetry breaking. We also emphasise that as very large volumes arise naturally in this model there is no fine-tuning problem in having $m_{SUSY} \sim \mathcal{O}(1\text{TeV})$.

Moduli stabilisation and the landscape have received much discussion recently. Technically, the landscape large numbers of $\mathcal{O}(10^{500})$ arise from the very many ways of stabilising moduli. It is clearly necessary to find a handle for dealing with such numbers. One developed approach is statistical ([39] and references thereto). Without explicitly constructing examples of vacua, this aims at framing and answering questions about what is and what is not possible. However, it may be very difficult to identify the right vacuum: it has been recently argued that the problem of finding a vacuum with the right cosmological constant is NP hard [40].

A more general point argued here is that in the context of the landscape the strong CP problem may serve as an *experimentum crucis*. Assuming that the solution to the strong CP problem is a Peccei-Quinn axion and that string theory is a correct description of nature, this is a solution that is extremely sensitive to the physics of moduli stabilisation. Requiring an axion to remain (essentially) massless while all other moduli are stabilised is a technically clean problem directly addressing the issue of vacuum selection. Indeed, as seen above imposing

this requirement directly rules out many scenarios of moduli stabilisation. The further condition $10^9\text{GeV} < f_a < 10^{12}\text{GeV}$ is even more constraining: we have described one approach to this above and it would be very interesting to analyse others.

Acknowledgements

I am funded by EPSRC and Trinity College, Cambridge. For conversations related to this work, I would like to thank M. Berg, C. Burgess, M. Dine, S. Kachru, F. Quevedo and P. Svrcek. I am also particularly grateful to S. Kachru and F.Quevedo for commenting on drafts of this manuscript. This work was initiated in the stimulating atmosphere of the Stanford Institute of Theoretical Physics, whose hospitality I gratefully acknowledge.

A Constraining Corrections to the Kähler Potential

Here I discuss how to constrain the form of corrections to the Kähler potential. The tree-level Kähler potential used for our toy model was

$$\mathcal{K} = -2\ln\left((T_1+\bar{T}_1)^{3/2} - (T_2+\bar{T}_2)^{3/2} - (T_3+\bar{T}_3)^{3/2}\right). \tag{82}$$

This geometry has one overall Kähler mode (T_1) and two blow-ups (T_2 and T_3). We specialise to our limit of interest $\tau_1 \gg \tau_2, \tau_3 > 1$, with $\tau_i = \text{Re}(T_i)$, where the resulting Kähler metric has the form (neglecting terms subleading in $\mathcal{V}$)

$$\mathcal{K}_{i\bar{j}} = \begin{pmatrix} \frac{3}{\mathcal{V}^{4/3}} & -\frac{9\sqrt{\tau_2}}{2\mathcal{V}^{5/3}} & -\frac{9\sqrt{\tau_3}}{2\mathcal{V}^{5/3}} \\ -\frac{9\sqrt{\tau_2}}{2\mathcal{V}^{5/3}} & \frac{3}{2\sqrt{2\tau_2}\mathcal{V}} & \frac{9\sqrt{\tau_2\tau_3}}{\mathcal{V}^2} \\ -\frac{9\sqrt{\tau_3}}{2\mathcal{V}^{5/3}} & \frac{9\sqrt{\tau_2\tau_3}}{\mathcal{V}^2} & \frac{3}{2\sqrt{2\tau_3}\mathcal{V}} \end{pmatrix}. \tag{83}$$

Any correction to the Kähler potential will also generate corrections to the Kähler metric (83). As such corrections are perturbative, they may arise either from α' effects or loop effects. On physical grounds, we expect that such corrections will be subdominant to the tree-level metric in the regime - large overall volume, large cycle volumes and weak coupling - where both worldsheet and quantum corrections ought to be least important.

We can apply this condition to restrict the form of potential corrections to $\mathcal{K}$. For example, consider the possible correction

$$\mathcal{K} + \delta\mathcal{K} = -2\ln(\mathcal{V}) + \frac{\epsilon\sqrt{2\tau_2}}{\mathcal{V}^\alpha}, \tag{84}$$

with $0 < \alpha < 1$. The $2\bar{2}$ component of the corrected Kähler metric is

$$\mathcal{K}_{2\bar{2}} + \delta\mathcal{K}_{2\bar{2}} = \frac{3}{2\sqrt{2\tau_2}\mathcal{V}} - \frac{\epsilon}{8\sqrt{2}\tau_2^{3/2}\mathcal{V}^\alpha}. \tag{85}$$

As $\alpha < 1$, the correction to the kinetic term would always dominate the tree-level term in the limit $\mathcal{V} \gg 1$. This seems implausible as a large-volume limit ought to make the correction less, rather than more, important.

A similar comment applies to a correction

$$\mathcal{K} + \delta\mathcal{K} = -2\ln(\mathcal{V}) + \frac{\epsilon\tau_2^2}{\mathcal{V}}, \tag{86}$$

which leads to

$$\mathcal{K}_{2\bar{2}} + \delta\mathcal{K}_{2\bar{2}} = \frac{3}{2\sqrt{2\tau_2}\mathcal{V}} + \frac{\epsilon}{4\mathcal{V}}. \tag{87}$$

In this case, as we take τ_2 large, the correction becomes increasingly dominant over the tree-level term. Given that large τ_2 reduces both the curvature of this cycle and the gauge coupling on any brane wrapping it, we would again expect exactly opposite behaviour to occur.

Finally, we could also consider the correction

$$\mathcal{K} + \delta\mathcal{K} = -2\ln(\mathcal{V}) + \frac{2\epsilon\tau_2}{\mathcal{V}^\alpha}, \tag{88}$$

with $0 < \alpha < 1$. In this case $\delta K_{2\bar{2}}$ is subleading to $\mathcal{K}_{2\bar{2}}$ in the classical limit. However, if we consider the $1\bar{2}$ component we now have

$$\mathcal{K}_{1\bar{2}} + \delta\mathcal{K}_{1\bar{2}} = -\frac{9\sqrt{\tau_2}}{2\mathcal{V}^{5/3}} + \frac{3\epsilon}{2\mathcal{V}^{\alpha+2/3}}, \tag{89}$$

and the correction again dominates in the large-volume limit.

The correction used in the body of the paper,

$$\mathcal{K} + \delta\mathcal{K} = -2\ln(\mathcal{V}) + \frac{\epsilon\sqrt{\tau_2}}{\mathcal{V}} + \frac{\epsilon\sqrt{\tau_3}}{\mathcal{V}}, \tag{90}$$

does not suffer from the above problems. As a correction to the Kähler metric, it gives

$$\mathcal{K}_{2\bar{2}} + \delta\mathcal{K}_{2\bar{2}} = \frac{3}{2\sqrt{2\tau_2}\mathcal{V}} - \frac{\epsilon}{16\tau_2^{3/2}\mathcal{V}}, \tag{91}$$

$$\mathcal{K}_{3\bar{3}} + \delta\mathcal{K}_{3\bar{3}} = \frac{3}{2\sqrt{2\tau_3}\mathcal{V}} - \frac{\epsilon}{16\tau_3^{3/2}\mathcal{V}}. \tag{92}$$

Such corrections are suppressed compared to the tree-level term by a factor τ^{-1}, i.e. g^2 of the field theory on the brane. Unlike those considered above, these

corrections are well-behaved (i.e. subdominant) in the classical limit, and there does not exist a 'bad' scaling limit.

We note that for the case where Kähler corrections have been computed [41], the correction does fall into this form. We only focus on the Kähler moduli dependence: the full expressions can be found in [41]. The correction gives

$$\mathcal{K} + \delta\mathcal{K} = -\ln(T_1 + \bar{T}_1) - \ln(T_2 + \bar{T}_2) - \ln(T_2 + \bar{T}_3) + \sum_{i=1}^{3} \frac{\epsilon_i}{T_i + \bar{T}_i}, \tag{93}$$

with for example

$$\mathcal{K}_{1\bar{1}} + \delta\mathcal{K}_{1\bar{1}} = \frac{1}{4\tau_1^2} + \frac{\epsilon_1}{4\tau_1^3}. \tag{94}$$

The loop-corrected Kähler metric is suppressed by a factor $\tau_1^{-1} = g^2$.

The point of this discussion is that the form of possible corrections to the Kähler potential can be heavily constrained by the reasonable requirement that in a classical (large volume, weak coupling) limit, corrections to the metric become increasingly subdominant to tree-level terms: simply because Kähler corrections are very hard to calculate does not make us entirely ignorant of their form. In particular, denoting the 'small', blow-up moduli by τ_i, these considerations exclude corrections of the form

$$\mathcal{K} + \delta K = -2\ln(\mathcal{V}) + \frac{f(\tau_i)}{\mathcal{V}^\alpha}, \tag{95}$$

with $\alpha < 1$, as in a classical, large volume limit there will be metric components whose correction dominates the tree-level term.

This motivates our use of the correction (90): it comes from a reasonable assumption about the form of the corrections to the Kähler metric, and is consistently subdominant in the classical limit.

References

[1] P. Candelas, G. T. Horowitz, A. Strominger and E. Witten, Nucl. Phys. B **258**, 46 (1985).

[2] V. Braun, Y. H. He, B. A. Ovrut and T. Pantev, arXiv:hep-th/0505041.

[3] R. Blumenhagen, M. Cvetic, P. Langacker and G. Shiu, arXiv:hep-th/0502005.

[4] G. Aldazabal, L. E. Ibanez, F. Quevedo and A. M. Uranga, JHEP **0008** (2000) 002 [arXiv:hep-th/0005067].

[5] H. Verlinde and M. Wijnholt, arXiv:hep-th/0508089.

[6] R. D. Peccei and H. R. Quinn, Phys. Rev. D **16**, 1791 (1977); R. D. Peccei and H. R. Quinn, Phys. Rev. Lett. **38**, 1440 (1977).

[7] M. Dine, arXiv:hep-ph/0011376.

[8] T. Banks, M. Dine and E. Gorbatov, JHEP **0408** (2004) 058 [arXiv:hep-th/0309170].

[9] J. F. Donoghue, Phys. Rev. D **69**, 106012 (2004) [Erratum-ibid. D **69**, 129901 (2004)] [arXiv:hep-th/0310203].

[10] V. Balasubramanian, P. Berglund, J. P. Conlon and F. Quevedo, JHEP **0503**, 007 (2005) [arXiv:hep-th/0502058].

[11] J. P. Conlon, F. Quevedo and K. Suruliz, JHEP **0508**, 007 (2005) [arXiv:hep-th/0505076].

[12] T. Banks and M. Dine, Nucl. Phys. B **505** (1997) 445 [arXiv:hep-th/9608197].

[13] T. Banks, M. Dine and M. Graesser, Phys. Rev. D **68** (2003) 075011 [arXiv:hep-ph/0210256].

[14] M. Endo, K. Hamaguchi and F. Takahashi, arXiv:hep-ph/0602061.

[15] S. Nakamura and M. Yamaguchi, arXiv:hep-ph/0602081.

[16] J. Polchinski, Phys. Rev. Lett. **75**, 4724 (1995) [arXiv:hep-th/9510017].

[17] T. W. Grimm and J. Louis, Nucl. Phys. B **699**, 387 (2004) [arXiv:hep-th/0403067].

[18] H. Jockers and J. Louis, Nucl. Phys. B **718**, 203 (2005) [arXiv:hep-th/0502059].

[19] T. Banks, D. B. Kaplan and A. E. Nelson, Phys. Rev. D **49**, 779 (1994) [arXiv:hep-ph/9308292].

[20] B. de Carlos, J. A. Casas, F. Quevedo and E. Roulet, Phys. Lett. B **318**, 447 (1993) [arXiv:hep-ph/9308325].

[21] S. B. Giddings, S. Kachru and J. Polchinski, Phys. Rev. D **66**, 106006 (2002) [arXiv:hep-th/0105097].

[22] S. Kachru, R. Kallosh, A. Linde and S. P. Trivedi, Phys. Rev. D **68**, 046005 (2003) [arXiv:hep-th/0301240].

[23] S. Gukov, C. Vafa and E. Witten, Nucl. Phys. B **584**, 69 (2000) [Erratum-ibid. B **608**, 477 (2001)] [arXiv:hep-th/9906070].

[24] E. Witten, Nucl. Phys. B **474**, 343 (1996) [arXiv:hep-th/9604030].

[25] F. Denef and M. R. Douglas, JHEP **0405** (2004) 072 [arXiv:hep-th/0404116].

[26] K. Becker, M. Becker, M. Haack and J. Louis, JHEP **0206** (2002) 060 [arXiv:hep-th/0204254].

[27] G. von Gersdorff and A. Hebecker, Phys. Lett. B **624** (2005) 270 [arXiv:hep-th/0507131].

[28] M. Berg, M. Haack and B. Kors, arXiv:hep-th/0508171.

[29] B. C. Allanach, F. Quevedo and K. Suruliz, arXiv:hep-ph/0512081.

[30] K. Choi, A. Falkowski, H. P. Nilles and M. Olechowski, Nucl. Phys. B **718**, 113 (2005) [arXiv:hep-th/0503216]; K. Choi, K. S. Jeong and K. i. Okumura, JHEP **0509** (2005) 039 [arXiv:hep-ph/0504037]; M. Endo, M. Yamaguchi and K. Yoshioka, Phys. Rev. D **72**, 015004 (2005) [arXiv:hep-ph/0504036].

[31] P. Breitenlohner and D. Z. Freedman, Annals Phys. **144** (1982) 249.

[32] O. DeWolfe, A. Giryavets, S. Kachru and W. Taylor, JHEP **0502** (2005) 037 [arXiv:hep-th/0411061].

[33] G. Villadoro and F. Zwirner, JHEP **0506** (2005) 047 [arXiv:hep-th/0503169]; G. Villadoro and F. Zwirner, arXiv:hep-th/0602120.

[34] O. DeWolfe, A. Giryavets, S. Kachru and W. Taylor, JHEP **0507** (2005) 066 [arXiv:hep-th/0505160].

[35] F. Denef, M. R. Douglas and B. Florea, JHEP **0406**, 034 (2004) [arXiv:hep-th/0404257].

[36] D. Lust, S. Reffert and S. Stieberger, Nucl. Phys. B **706**, 3 (2005) [arXiv:hep-th/0406092].

[37] P. G. Camara, L. E. Ibanez and A. M. Uranga, Nucl. Phys. B **708**, 268 (2005) [arXiv:hep-th/0408036].

[38] C. P. Burgess, L. E. Ibanez and F. Quevedo, Phys. Lett. B **447** (1999) 257 [arXiv:hep-ph/9810535].

[39] M. R. Douglas, JHEP **0305** (2003) 046 [arXiv:hep-th/0303194].

[40] F. Denef and M. R. Douglas, arXiv:hep-th/0602072.

[41] M. Berg, M. Haack and B. Kors, JHEP **0511** (2005) 030 [arXiv:hep-th/0508043].